

User Interface Issues in Pervasive Computing

Presented at NIST Pervasive Computing 2001 Conference

Ponani S. Gopalakrishnan

Manager, Conversational Systems

IBM T.J. Watson Research Center

P.O. Box 218

Yorktown Heights, NY 10598

psg@us.ibm.com

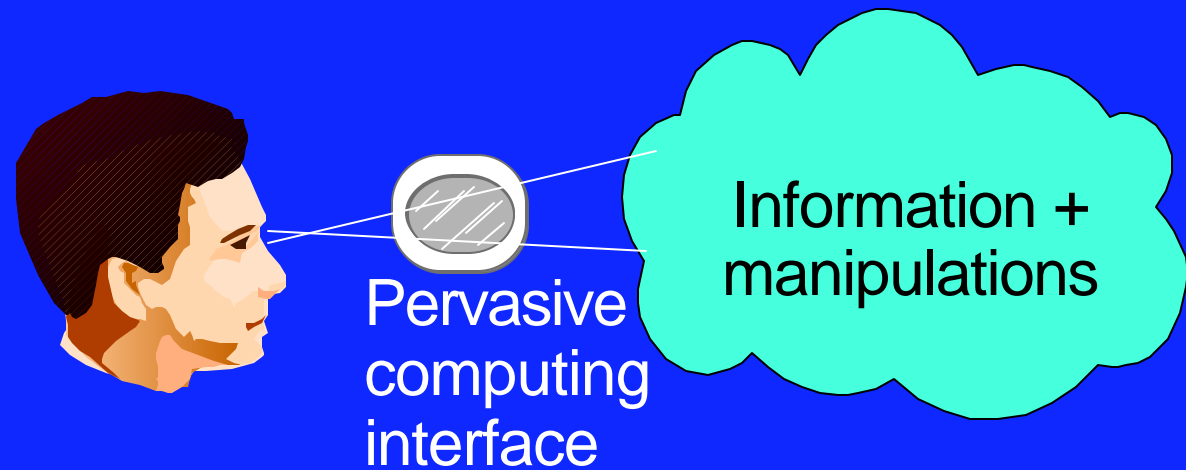


Outline

- Pervasive Computing Interfaces
 - Complexity of interaction
 - Interface capabilities
- Desired characteristics
- Some examples
- Lessons learned
- Ongoing and future work

Pervasive Computing Interface Characteristics

- Complexity of interaction



Focus on more complex interactions

- Interface constrained by available sensors and actuators
 - Keyboard
 - Stylus
 - Audio input / output
 - Video input
 - Touch screen
 - ...



Interface Requirements

- Support Mobility
- Eyes-busy hands-busy
- Small form factor
- Multimodal
- Personalization
- Assistive
- Consistency across different instances
- Natural

What is Available Today

- Stylus and touchscreen
- Handwriting / Graffiti
- Telephone keypad
- Voice
- Limited video input



Prototypes

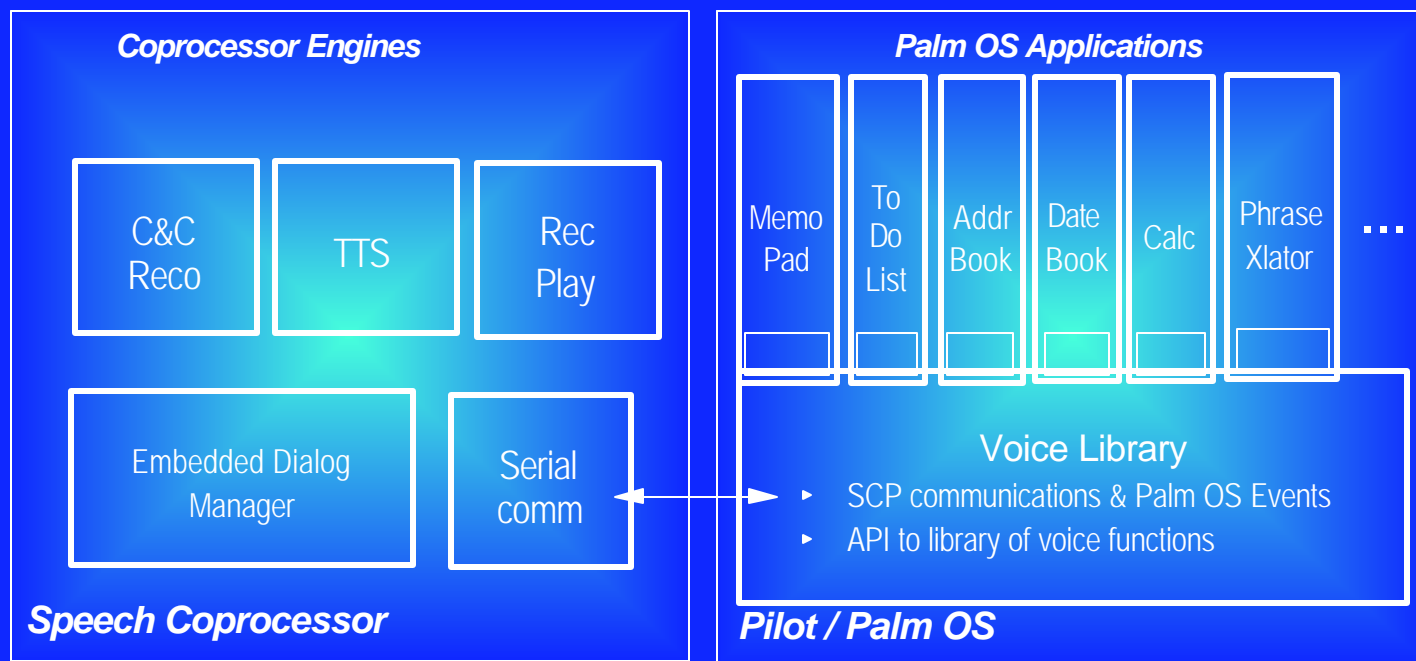
- Two research projects at IBM
 - Personal Speech Assistant
 - Conversational Assistant for Email
- Domains
 - PIM, Email
- UI capabilities
 - Voice input and output combined with other modalities

Personal Speech Assistant

- ◆ Speech-enabled PIM applications
 - *Datebook, address book, to-do list, phrase translator*
- ◆ GUI functions unchanged
- ◆ Helper applications on PC
- ◆ Hot sync to install new applications
- ◆ User-configurable commands



Personal Speech Assistant



◆ Speech recognition

- *IBM ViaVoice Embedded*
 - *Medium vocabulary (500 words), dynamic vocabulary size, far-field microphone*

◆ Text to speech

- *Full-function synthesis (US English), user-selectable voice characteristics, playback stored waveforms*

◆ Deferred recognition

- *Interface to IBM ViaVoice Dictation (PC) for off-line transcription*

What more can we do?

- Point and speak
- Rich GUI - full screen, keyboard
- Conversational

- *Donatello* - a conversational (multimodal) assistant
 - PC based
 - Email, calendar, address book functions
 - Rich GUI
 - Conversational natural language interactions
 - Inferencing and dialog management across modalities
 - Access from desktop or telephone
 - System "personality" and context preserved across different access methods
 - User-centered design

Example of Mixed-Initiative Dialog

U: Put all the messages from Jeff about proposal into a folder called proposal.

S: Do you mean Jeff Smith or Jeff Houston?

S: Excuse me, you have one new message.

U: Who is it from?

S: The author of the message is Susan McKay.

U: What's the subject?

S: The subject is let's move the meeting on September 12th to afternoon around 3

U: What day is September 12th?

S: That is a Friday.

U: Find the meeting with Susan on that day.

U: Move the meeting to 3 o'clock

U: Repeat that

S: Do you mean Jeff Smith or Jeff Houston?

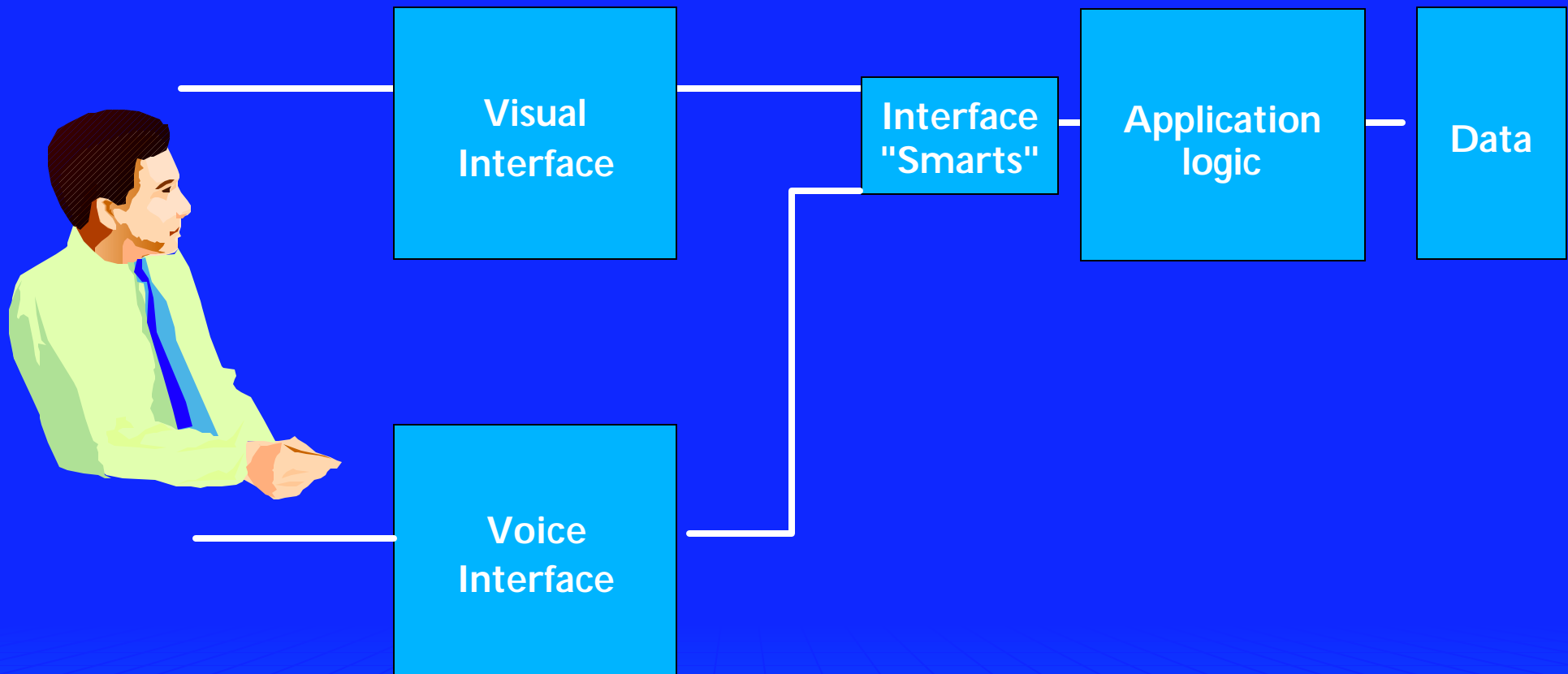
U: I mean Jeff Houston.

S: The folder called proposal does not exist. Would you like me to create it?

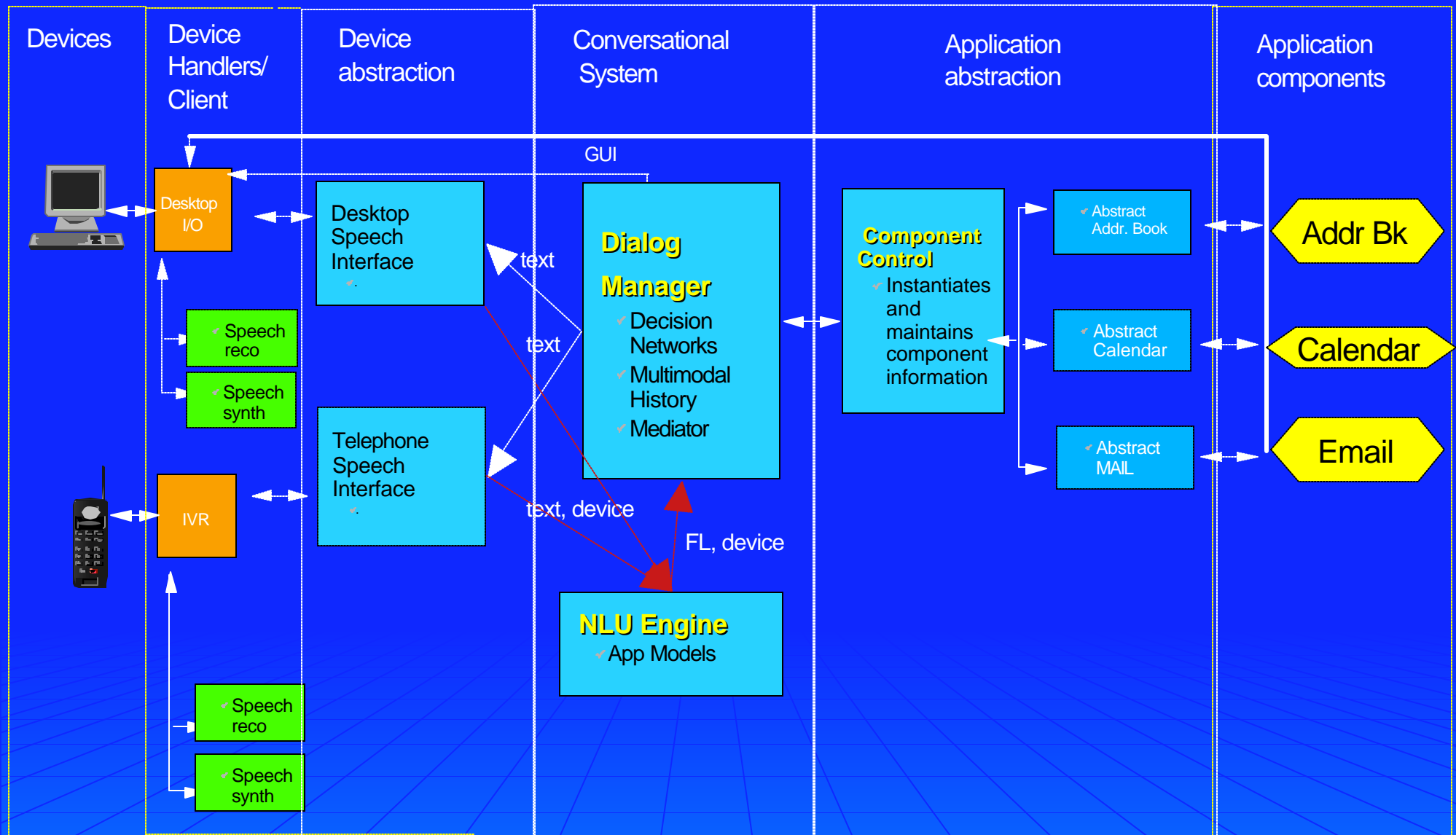
U: Yes

S: This may take a few seconds... done

Architecture



Architecture



Example Sentences From Data Collection

What day is April 12

Put all the messages from Jeffrey about proposal in a folder called proposal

Go to the bank at two, deposit checks, pick up dog food on the way home

Can you give me the names of the people who attended the staff meeting last Tuesday

Can you get a hold of Debra who visited me last Tuesday and see if she's able to come back on Friday.

I want to take half day vacation on February tenth so please ah can you cancel my all appointment from on February tenth from one P.M. to five P.M.

Schedule one hour with my manager for March one I have to meet with him and inform him about the antenna project in Hawthorne.

What did we learn ?

■ Strengths

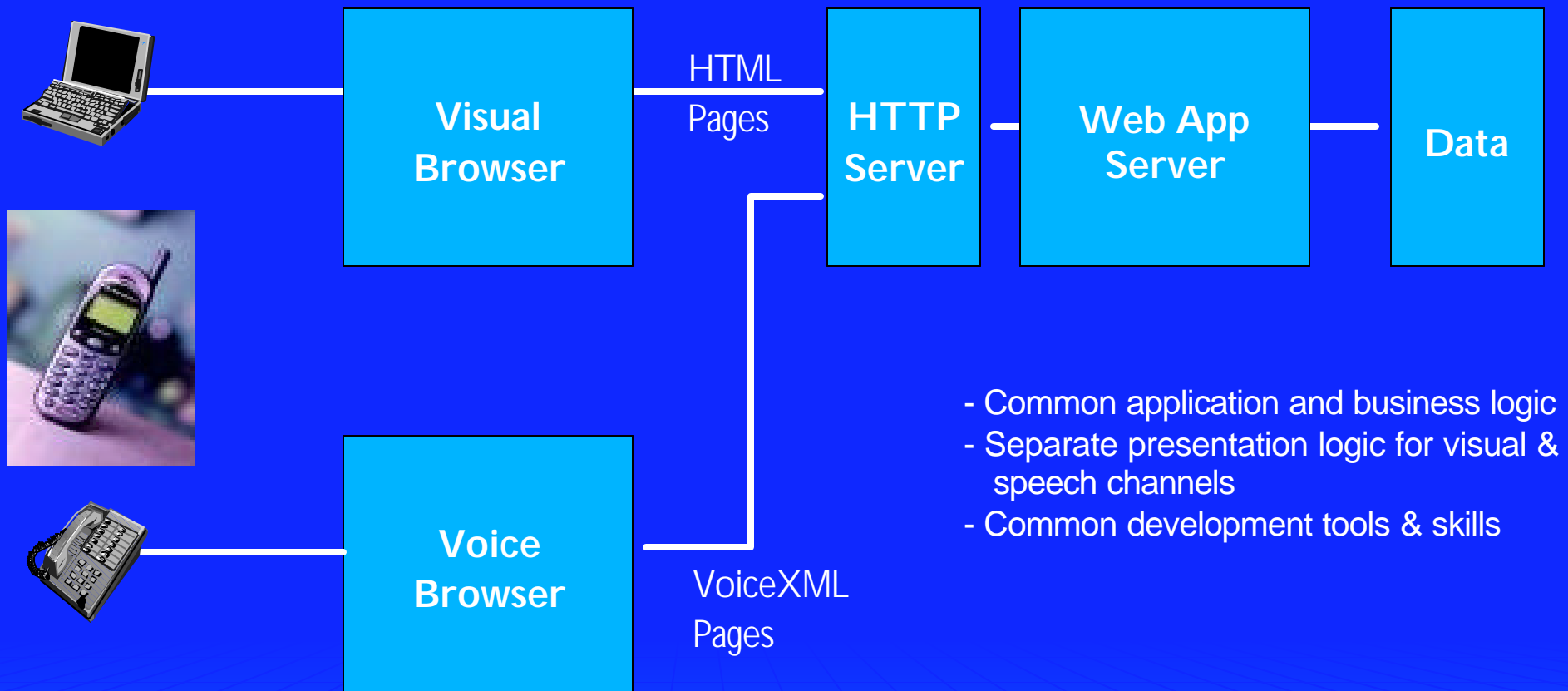
- Easier, natural information interaction
 - Searching through large amounts of data
 - Manipulating applications not in GUI focus
 - Manipulating multiple applications
- Multimodal user interaction
 - GUI action may be preferred for visible items
- Multimodal dialog management with reference resolution
- Multi-client user interaction
 - Seamless interaction across multiple devices

■ Enhancements needed

- Personalization
- Context dependent help (both user and system initiated)
 - Making user aware of state of discourse
- More system initiative
- Reminder and summarization features
- Natural language generation
- Transaction accuracy

Multimodal Browsing

■ Desired architecture



- Common application and business logic
- Separate presentation logic for visual & speech channels
- Common development tools & skills

Voice Access

Voice Browser

- Manages audio, call control, recognition, text-to-speech, dialog with user

Summary

- Renewed emphasis on user interface design
- Multimodal interfaces show promise
- UI Infrastructure needs to exist
- Focus on user centered design
- Many challenges still remain